Звіт з лабораторної роботи №1

З курсу «Комп’ютерні системи»

Студентки 3 курсу КІ-СА

Уваренко Олени

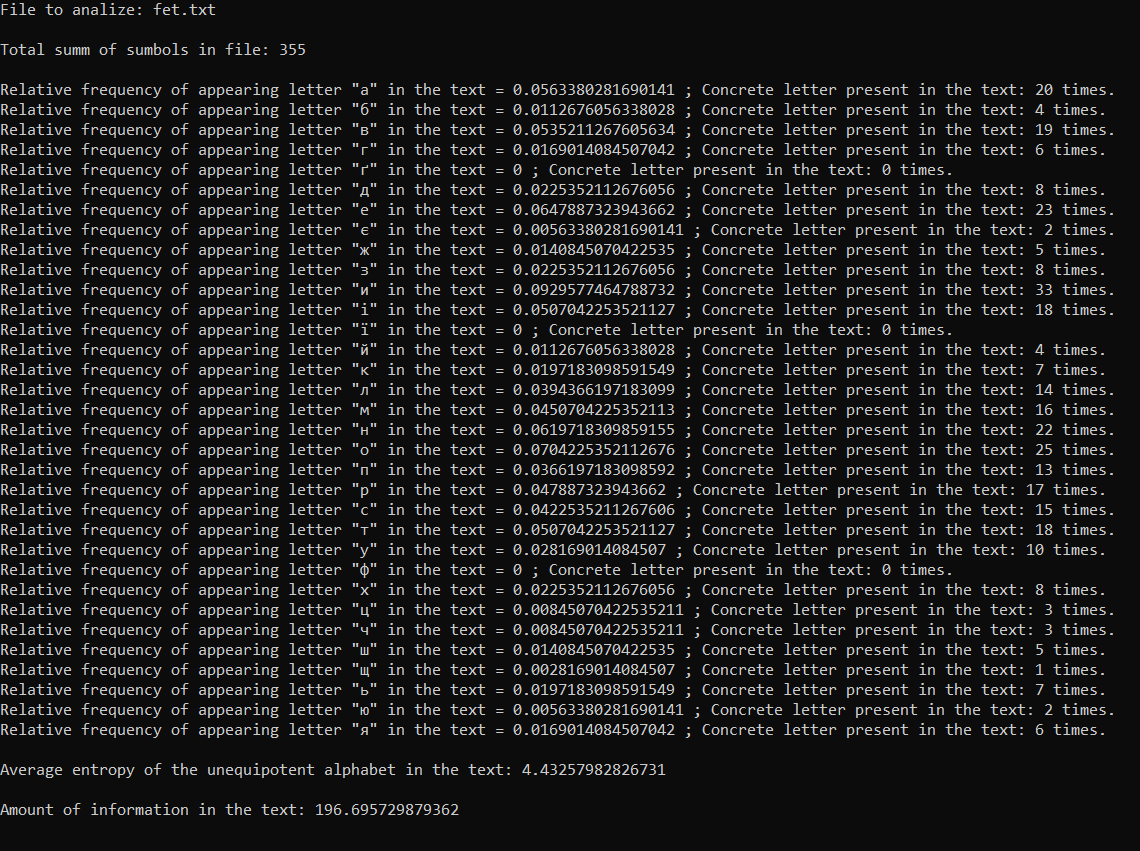
1. Оберіть 3 текстових файла різного тематичного та лінгвістичного спрямування (наприклад, вірш Тараса Шевченка “Мені тринадцятий минало”, “Казка про рєпку” Леся Подерв'янського та специфікацію інерфейсу PCI)

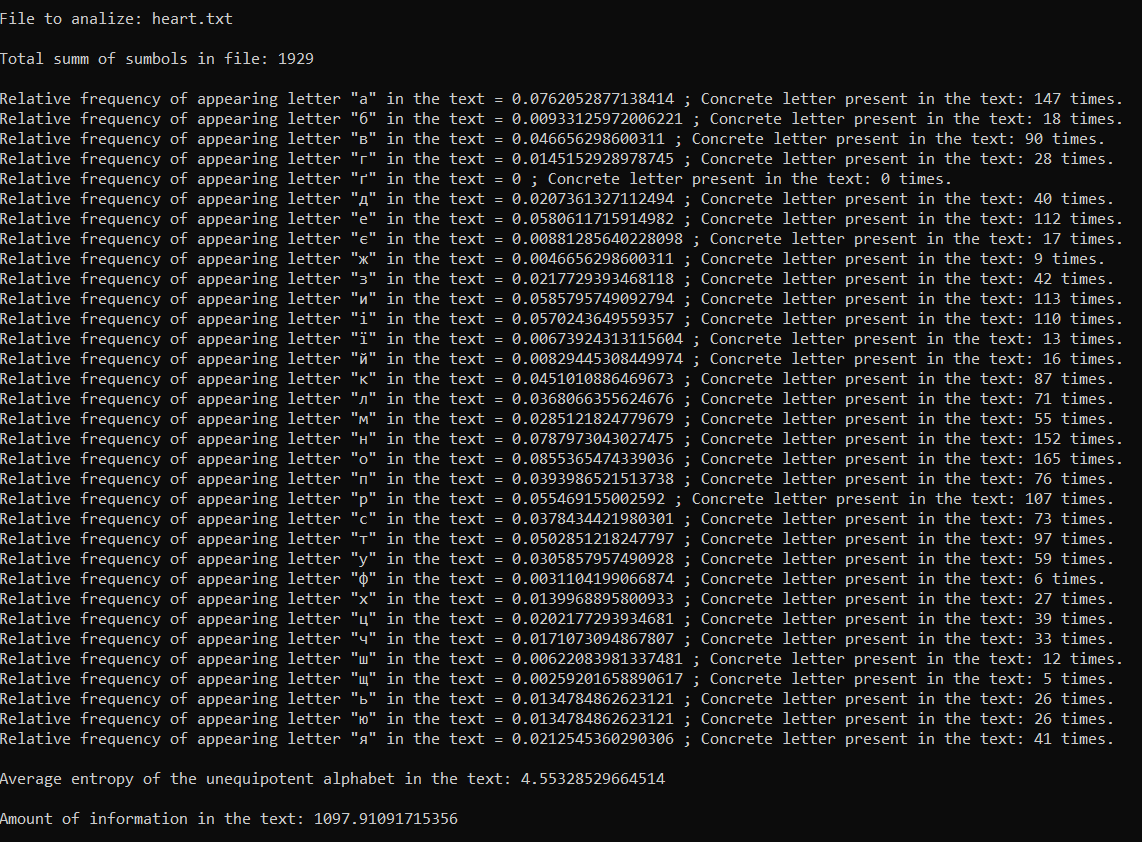
Мої файли:

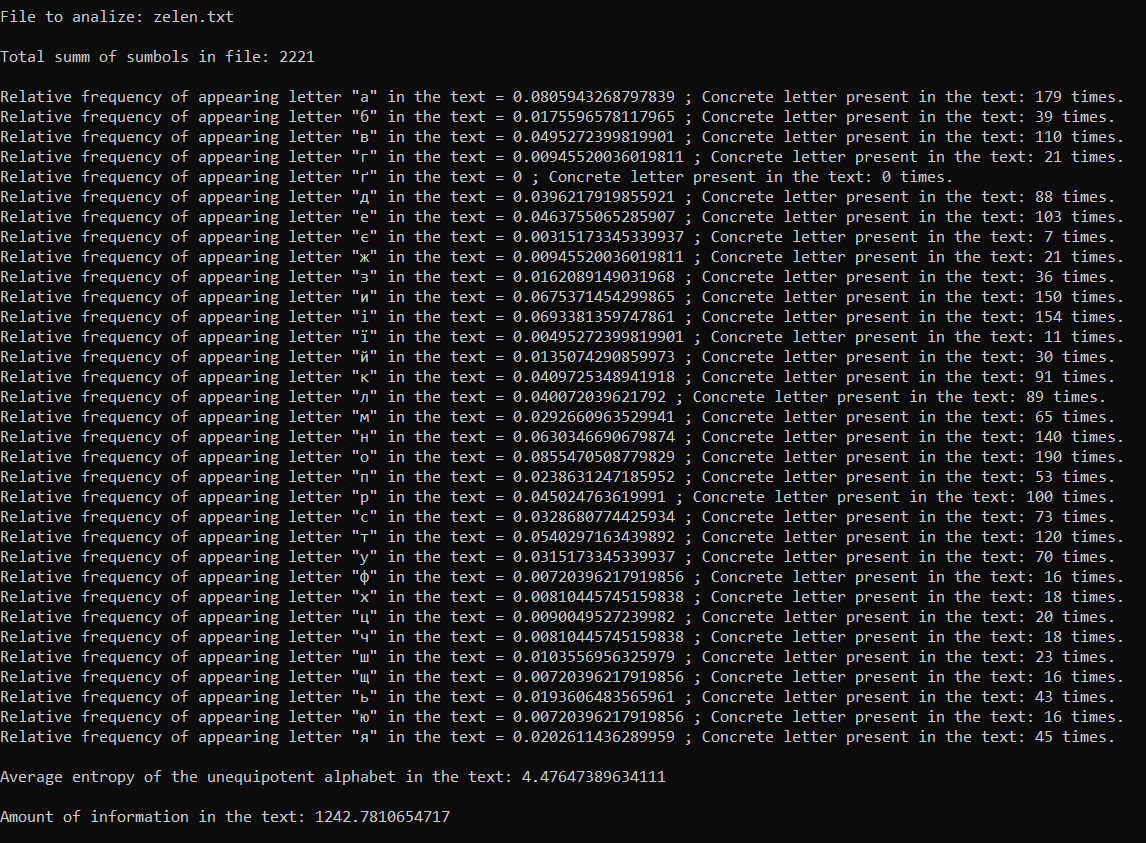
* + 1. Стаття «Події поза планами кандидатів у президенти: Зеленський, Порошенко»
    2. «Убога мова в нас» А.Фета
    3. Класифікація штучних клапанів для серця

1. Створіть програму (будь-якою зручною для вас мовою), яка в якості вхідних даних приймає текстовий файл, та аналізуючи його вміст:
   * 1. обраховує частоти (імовірності) появи символів в тексті
     2. обраховує середню ентропію алфавіту для даного тексту
     3. виходячи з ентропії визначає кількість інформації та порівнює її з розмірами файлів
     4. виводить на екран значення частот, ентропії та кількості інформації







1. Проведіть стиснення кожного вхідного файлу за допомогою 5 різних алгоритмів стиснення (zip, rar, gzip, bzip2, xz, або будь-які інші на ваш вибір, можна використовувати готові програмні засоби для стиснення).
2. Порівняйте результуючі обсяги архівів з обчисленою кількістю інформації та наведіть у звіті висновки щодо кореляції цих величин для обраних вами файлів (яка відмінність, що вийшло більше і чому)

Кількість інформації - мінімальна кількість біт, необхідна для кодування всіх можливих значень повідомлення. Вона буде меншою за розмір наший файлів.

З розповсюджених визначень, можна навести таке:

Кількість інформаціі  можна розгядати як міру зменшення невизначенності знання

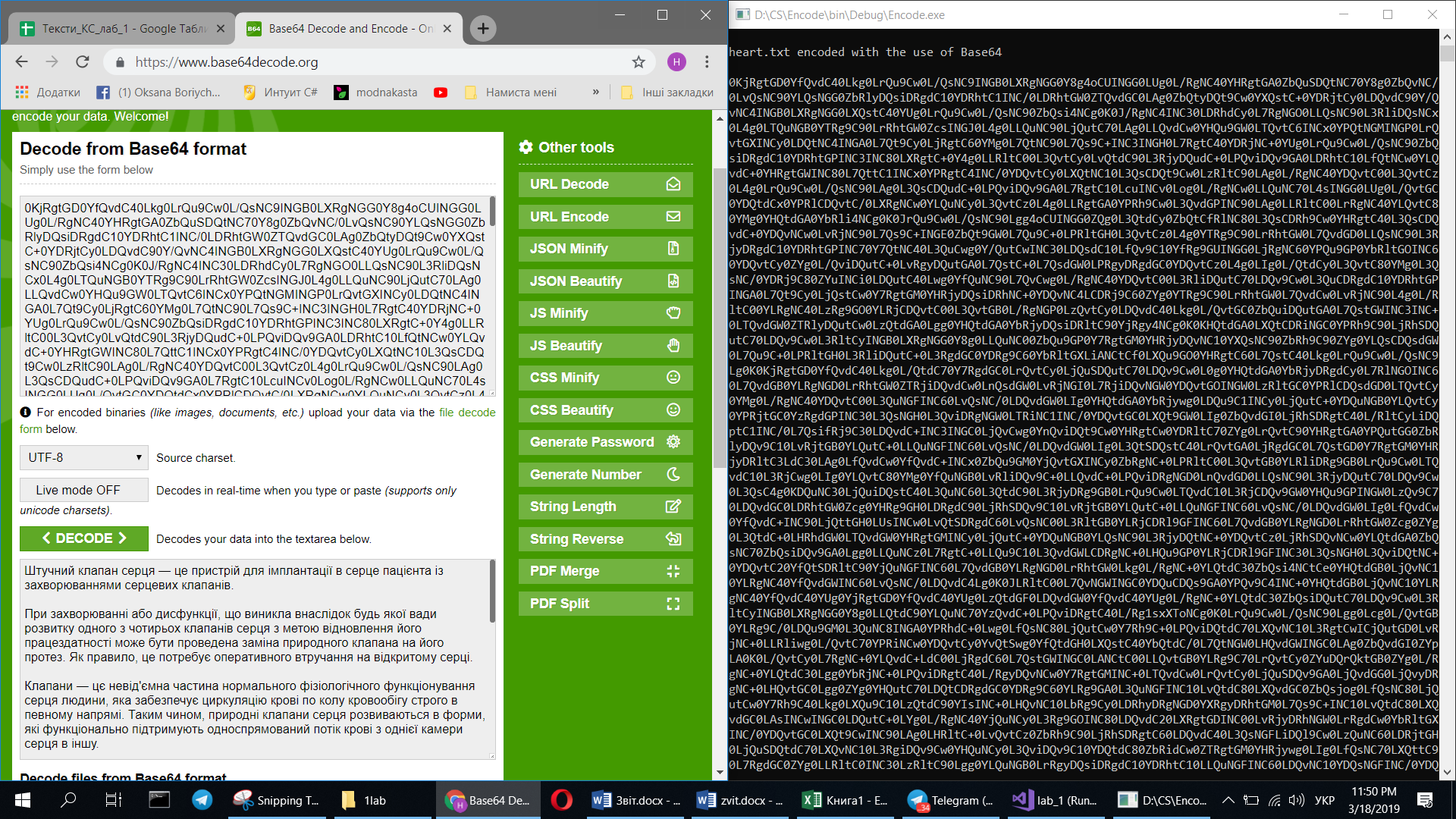
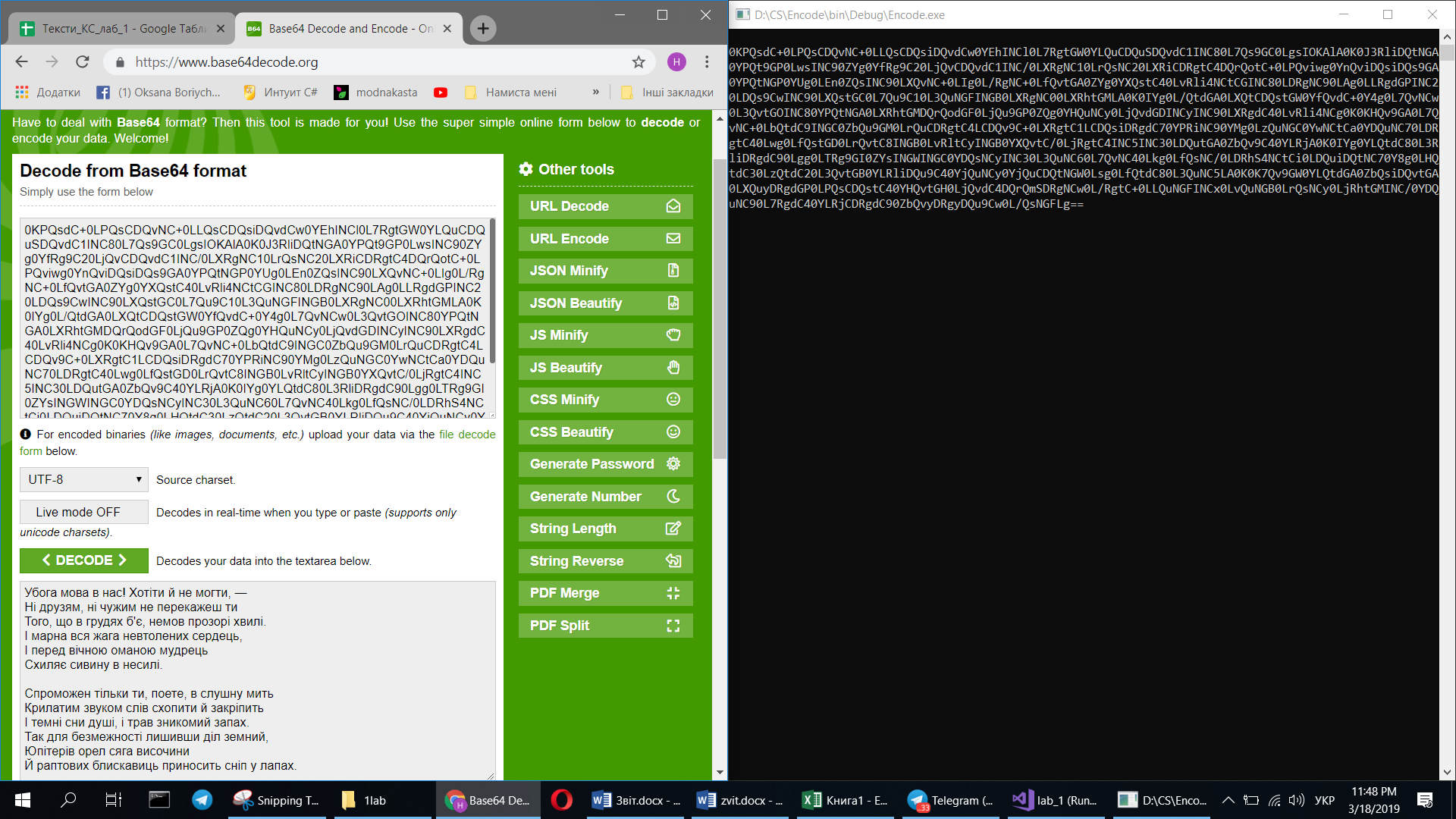
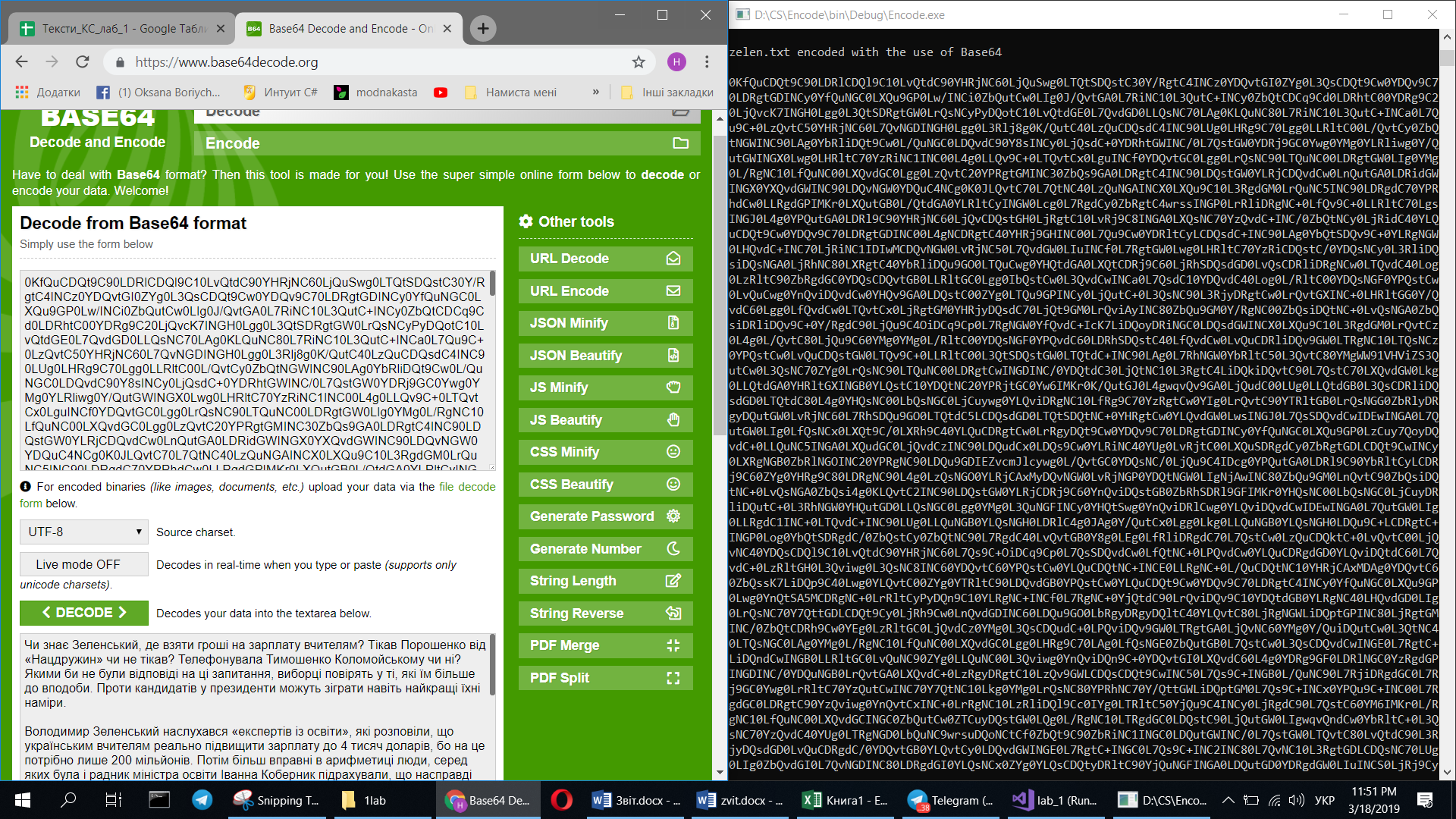
При отриманні інформаційних поідомень.

|  |  |  |  |
| --- | --- | --- | --- |
| Файл | fet.txt, bytes | heart.txt, bytes | zelen.txt, bytes |
| rar | 524 | 1,527 | 1,994 |
| tar | 2,560 | 6,144 | 6,656 |
| bz2 | 383 | 1,201 | 1,601 |
| zip | 577 | 1,579 | 2,051 |
| xz | 508 | 1,452 | 1,880 |
| ORIGINAL | 817 | 4,222 | 4,982 |
| Кількість інформації | 196 | 1097 | 1242 |

2. Дослідження способів кодування інформації на прикладі Base64

1. Ознайомтесь зі стандартом [RFC4648](https://tools.ietf.org/html/rfc4648)
2. Для практичного засвоєння методу кодування, створіть програму, що кодує довільний файл в Base64 (шляхом реалізації алгоритму вручну, а не виклику бібліотечної функції)

перевірте коректність роботи програми, порівнявши результат з існуючими програмними засобами (наприклад, openssl enc -base64)

1. Закодуйте в Base64 обрані вами текстові файли
2. 
   * 1. Обрахуйте кількість інформації в base64-закодованому варіанті файлу
     2. Порівняйте отримане значення з кількістю інформації вихідного файлу
     3. Зробіть висновки з отриманого результату
3. Закодуйте в Base64 стиснені кращим з алгоритмів текстові файли

1.Обрахуйте кількість інформації в base64-закодованому варіанті   
стисненого файлу

2. Порівняйте отримане значення з кількістю інформації вихідного   
файлу та base64-закодованого файлу

3. Зробіть висновки з отриманого результату

**Посилання на код програм:**

<https://github.com/helenuvarenko/Computer-Systems/tree/master/1lab>